УДК 378.146:51:681.3

КОМПЬЮТЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ КАЧЕСТВА ИНЖЕНЕРНОГО ОБРАЗОВАНИЯ. ВХОДНОЙ КОНТРОЛЬ МАТЕМАТИЧЕСКИХ ЗНАНИЙ

В.П. Арефьев, А.А. Михальчук, Н.Н. Кулебакина

Томский политехнический университет
E-mail: vpa@ido.tpu.edu.ru

Проведен компьютерный статистический сравнительный анализ результатов вступительных испытаний по математике с результатами традиционного входного контроля математических знаний в зависимости от разных форм испытания и обучения в вузе, а также сравнение результатов входного контроля с тестированием, использующим компьютерные технологии. Сделан вывод о статистически значимых различиях этих результатов. Обсуждаются причины выявленных существенных различий. Проведен компьютерный хронологический анализ и прогнозирование результатов входного контроля.

Введение

Главной проблемой современного инженерного образования является проблема его качества [1]. В связи с подписанием Россией Болонского соглашения одной из важных задач ближайшего времени станет разработка и использование методик оценки качества образования, сопоставимых с общеевропейскими. В частности, необходимо совершенствование и оптимизация приема в вузы, создание эффективной системы контроля качества знаний (например, тестирование) [2, 3].

В Томском политехническом университете (ТПУ) наряду с вступительными испытаниями (**ВИ**) по математике в разных формах испытаний (ФИ): ЕГЭ, централизованное тестирование (ЦТ), олимпиада (О), вступительный экзамен (ЭКЗ), проводится входной контроль (**ВК**) математических знаний школьной программы на основе аудиторной контрольной работы с проверкой ее преподавателями. Задание **ВК** содержит 6 задач средней сложности (типа группы В в билетах ЕГЭ). Результаты **ВК** (по 5-бальной шкале) за последние 10 лет представлены в табл. 1.

Таблица 1. *Средний балл m результатов **ВК** по 5-бальной шкале*

Год	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
m	1,906	1,838	1,788	1,770	2,195	2,250	2,308	2,250	2,381	2,216

Аналогичный **ВК** проводится и в других вузах России, например, в Кузбасском государственном техническом университете и Кемеровском государственном университете с 1971 г. [4].

В связи с этим представляет интерес сравнение результатов выше указанного **ВИ** с соответствующими результатами **ВК**.

Статистический метод

Все числовые результаты (например, **ВИ** по 100-бальной шкале и **ВК** по 12-бальной шкале) приведены к единой 5-бальной шкале (делением результата на соответствующий максимальный результат и умножением на пять). Созданная таким образом в MS Excel база данных использовалась далее в пакете Statistica 6.0 для статистического анализа данных: вычисления числовых характеристик (среднее $\boldsymbol{m}$, стандартное отклонение $\boldsymbol{\sigma}$, асимметрия $\boldsymbol{A}$ и эксцесс $\boldsymbol{E}$), применения параметрического $\boldsymbol{t}$-критерия Стьюдента для сравнения средних двух независимых нормальных выборок (для проверки нормальности распределения использованы оценки асимметрии $\boldsymbol{A}$ и эксцесса $\boldsymbol{E}$, χ^2 – критерий Пирсона и критерий Колмогорова-Смирнова), а также непараметрических критериев: критерий серий Вальда-Вольфовица, $\boldsymbol{U}$– критерий Манна-Уитни, критерий Колмогорова-Смирнова [5]. Применялся метод экспоненциального сглаживания и прогнозирования временного ряда m (табл. 1).

Результаты и их анализ

Во **ВК** 2004 г. участвовали 1175 студентов первого курса ТПУ: институт геологии и нефтегазового дела (ИГНД), электротехнический институт (ЭЛТИ); физико-технический (ФТФ), электрофизический (ЭФФ), машиностроительный (МСФ), химико-технологический (ХТФ), теплоэнергетический (ТЭФ),

автоматики и вычислительной техники (АВТФ), естественных наук и математики (ЕНМФ) факультеты. Данные по **ВИ** брались из ведомостей по **ВК** со слов студентов. Т.о., объем выборки результатов **ВИ** составил 1133 студента. При этом 42 «забывчивых» студента так и остались неопределенными в связи с отказом центральной приемной комиссии ТПУ уточнить эту информацию. Числовые характеристики выборок **ВИ** и **ВК** приведены в табл. 2.

Средний балл во **ВИ** составил ≈3,45 по 5-бальной шкале, в то время как во **ВК** только ≈2,22, то есть ниже более чем на балл. Можно заметить, что средний балл во **ВК** для студентов-заочников 1-го курса Беловского филиала ТПУ в 2004 г. ≈1,05 (выборка **ВКБл** в табл. 2), то есть еще ниже более чем на балл. Сравнительное распределение выборок **ВИ** и **ВК** по равномерной 5-бальной шкале отображено на рис. 1.

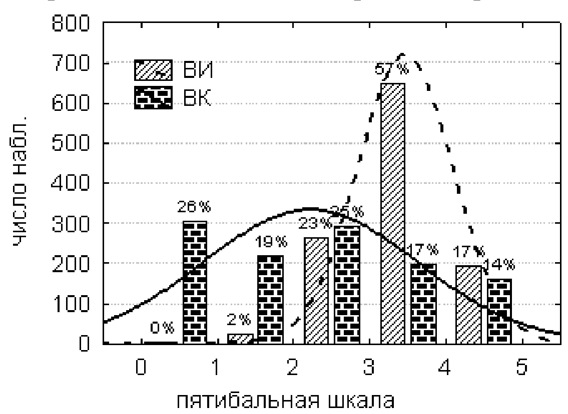


Рис. 1. *Составная гистограмма результатов **ВИ** и **ВК** в 2004 г*

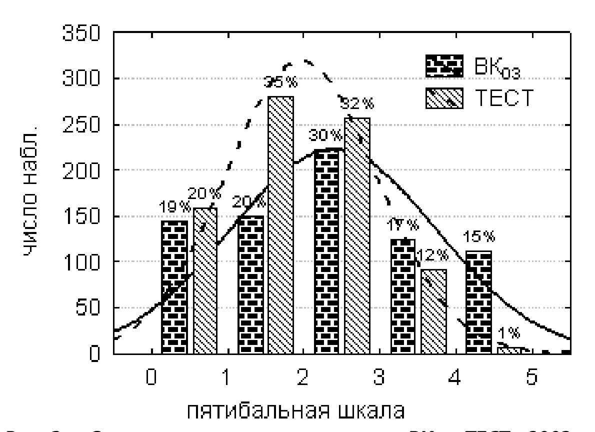


Рис. 2. *Составная гистограмма результатов $\textbf{ВК}_{03}$ и **ТЕСТ** в 2003 г.*

Над столбцами указаны значения относительных частот в процентах. Как следует из рис. 1, выборка результатов **ВИ** является более компактной (стандартное отклонение σ≈0,63), асимметричной вправо (коэффициент асимметрии A≈–0,52<0), остро вершинной (коэффициент эксцесса E≈0,28>0) и соответствует скорее нормальному закону распределения, в то время как выборка результатов **ВК** является более распыленной (стандартное отклонение σ≈1,41), асимметричной влево (коэффициент асимметрии A≈0,13>0), туповершинной (коэффициент эксцесса E≈–0,97<0) и соответствует скорее равномерному закону распределения.

Для оценки значимости различий в выборках результатов **ВИ** и **ВК** использовались критерии сравнения средних двух независимых выборок. В связи с проверкой применения t – критерия Стьюдента для визуальной оценки сходства наблюдаемых распределений (гистограммы) с теоретическим распределением по нормальному закону (соответствующие кривые) использована равномерная 5-бальная шкала. Проверка нормальности распределения выборок с помощью коэффициентов асимметрии A и эксцесса E (отношения A и E по модулю к соответствующим стандартным ошибкам не должны превышать числа 3) показывает существенное отклонение от нормальности в случае **ВИ** по асимметрии (0,52/0,073≈7,12>3) и **ВК** по эксцессу (0,97/0,143≈7,46>3) в соответствии с табл. 2. Проверка нормальности распределения выборок с помощью χ^2-критерия Пирсона и критерия Колмогорова-Смирнова дает высоко значимые (уровень значимости α<0,001) отличия распределений выборок от нормального закона. В связи с нарушением условия нормальности распределений выборок далее применялись непараметрические критерии сравнения средних двух независимых выборок, откуда следует вывод о высоко значимом различии в средних баллах 3,45 для **ВИ** и 2,22 для **ВК**.

В настоящее время в ТПУ проводится эксперимент по введению системы независимой оценки качества обучения в тестовой форме с использованием компьютерной технологии [6, 7], в рамках которого тестированием в форме, аналогичной билетам ЕГЭ (в объеме групп заданий А и В, без группы С), с последующей машинной проверкой оценивались остаточные знания школьной программы (**ТЕСТ**). К сожалению, в 2004 г. **ТЕСТ** не проводилось. Поэтому в данной работе проведено сравнение $\textbf{ВК}_{03}$ и **ТЕСТ** по результатам 2003 г. Сравнение $\textbf{ВК}_{03}$ и **ТЕСТ** проведено по аналогии со сравнением выборок **ВИ** и **ВК**. Числовые характеристики выборок $\textbf{ВК}_{03}$ и **ТЕСТ** приведены в табл. 2. Сравнительное распределение выборок $\textbf{ВК}_{03}$ и **ТЕСТ** по равномерной 5-бальной шкале отображено на рис. 2. Статистические критерии дали высоко значимые (α<<0,001) различия в средних баллах результатов входного контроля 1,96 для **ТЕСТ** и 2,38 для $\textbf{ВК}_{03}$. Таким образом, **ТЕСТ** определяет уровень знаний студентов ниже, чем ВК. Предложенные тестовые механизмы оценки знаний являются более грубыми, чем существующие традиционные. По аналогии с билетами ЕГЭ билеты **ТЕСТ** содержат только задания на «угадывание» (группа А) и задания, правильность решения которых проверяются только по конечному числовому результату (группа В).

Ниже результаты **ВИ** и **ВК** по 2004 г. рассмотрены по факультетам. Значения основных числовых характеристик выборок **ВИ** и **ВК** по факультетам приведены в табл. 3. Значения объема N выборок **ВИ** и **ВК** по факультетам изменяются в широком диапазоне от 51 до 214 (см. также рис. 4). Значения m и σ выборок **ВИ** и **ВК** по факультетам отличаются от соответствующих m и σ выборок **ВИ** и **ВК** по ТПУ

(табл. 2) на величину от нескольких сотых до нескольких десятых долей по 5-бальной шкале. При этом, как показывает применение непараметрических критериев сравнения средних выборок по факультету и по ТПУ, различие даже в одну десятую долю по 5-бальной шкале приводит к статистически значимым отличиям (уровень значимости α<0,05).

Таблица 2. *Числовые характеристики выборок **ВИ, ВК, ВК$_{03}$, ТЕСТ и ВКБл***

Выборка	Объем N	Среднее m	Станд. отклонение σ	Асимметрия A	Станд. ошибка асимметрии	Эксцесс E	Станд. ошибка эксцесса
ВИ	1133	3,45	0,63	**−0,52**	**0,073**	0,28	0,145
ВК	1175	2,22	1,41	0,13	0,071	**−0,97**	**0,143**
ВК$_{03}$	754	2,38	1,35	0,028	0,089	**−0,811**	**0,178**
ТЕСТ	796	1,96	0,99	-0,232	0,087	−0,447	0,173
ВКБл	129	1.05	0,84	0,631	0,213	0,051	0,423

Таблица 3. *Числовые характеристики выборок **ВИ** и **ВК** по факультетам*

Факультет	ФТФ	ЭФФ	ХТФ	ЕНМФ	ЭЛТИ	ИГНД	МСФ	ТЭФ	АВТФ
$N_{ВИ}$	214	169	144	51	105	162	64	54	170
$N_{ВК}$	217	174	147	51	115	175	64	55	177
$m_{ВИ}$	3,37	3,49	3,36	3,40	3,31	3,39	3,33	3,55	3,76
$m_{ВК}$	2,32	2,05	2,07	2,36	1,92	2,06	1,91	2,51	2,70
$\sigma_{ВИ}$	0,56	0,61	0,58	0,76	0,76	0,57	0,77	0,57	0,55
$\sigma_{ВК}$	1,37	1,32	1,36	1,67	1,31	1,41	1,44	1,43	1,39

Значения средних баллов результатов **ВИ** и **ВК** по факультетам с указанием 95 % доверительного интервала приведены на рис. 3. Ширина доверительного интервала, определяющая погрешность измерения среднего, пропорциональна σ и обратно пропорциональна $\sqrt{N}$ [8]. Таким образом, более широкие доверительные интервалы m для ЕНМФ, МСФ и ТЭФ объясняются в большей мере сравнительно малыми объемами этих выборок N (рис. 4).

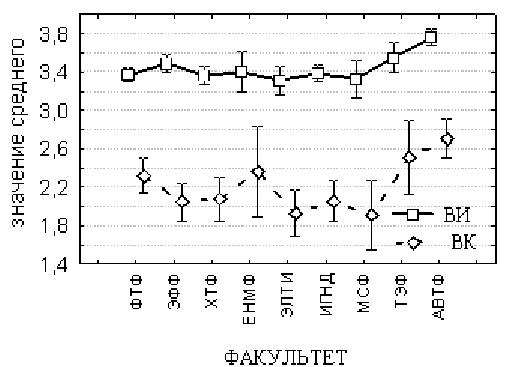


Рис. 3. *Линейные графики средних баллов с 95 % доверительными интервалами результатов **ВИ** и **ВК** по факультетам*

Сравнительной характеристикой различия выборок **ВИ** и **ВК** может быть распределение выборки **ВИ** – **ВК** (рис. 5) разностей результатов выборок **ВИ** и **ВК**.

Как следует из рис. 5, 82,5 %, т.е. ≈4/5 студентов из 1133 не подтвердили на **ВК** уровня своих знаний, отраженного в **ВИ**. При этом, например, 7,8 % студентов имеют результаты на **ВК** ниже соответствующих результатов в **ВИ** более чем на три балла по 5-бальной шкале.

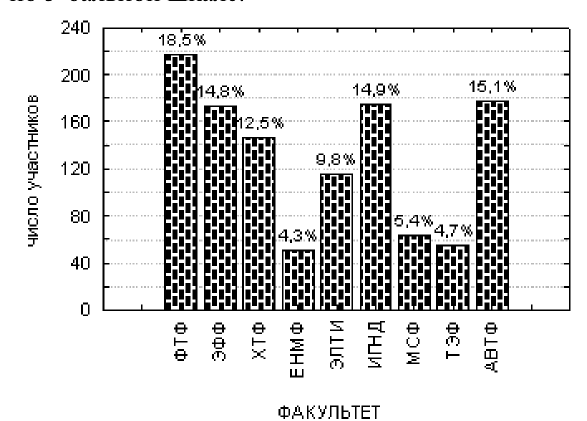


Рис. 4. *Гистограмма числа участников **ВК** по факультетам*

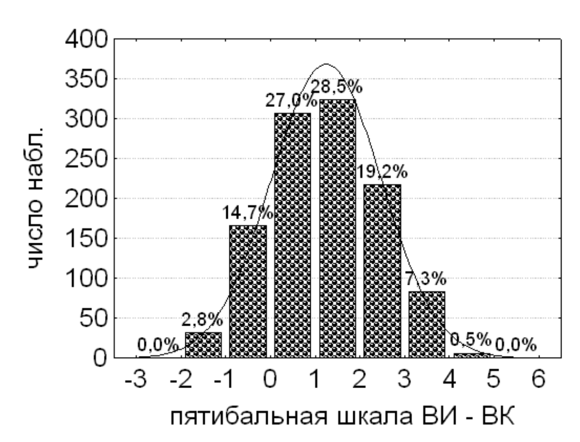


Рис. 5. *Гистограмма распределения разности значений результатов **ВИ** и **ВК***

Представляет интерес исследование результатов **ВИ** и **ВК** в зависимости от формы испытания (ФИ) (среди участвовавших во **ВК** 73,9 % , т.е. ≈3/4 поступили по результатам ЕГЭ, 10,2 % – по ЦТ, 8,5 % – по ЭКЗ и 7,4 % – по О), отраженное на рис. 6 и 7, где для простоты восприятия использована неравномерная 5-бальная шкала: [0; 2,5] – «неуд»; (2,5;5] – «уд + хор + отл». Из сравнения средних баллов следует, что наиболее низкие средние баллы как в **ВИ**, так и на **ВК** в следующих ФИ: ЭКЗ и ЦТ. При этом средние баллы в **ВИ** и во **ВК** отличаются почти в два раза. Сравнение результатов **ВИ** и **ВК** в зависимости от ФИ отражает ситуацию в целом: различия результатов **ВИ** и **ВК** высоко значимы в каждой **ФИ**.

На рис. 8 и 9 отражены аналогичные результаты **ВИ** и **ВК** в зависимости от формы обучения (ФО) среди участвовавших во **ВК** 67,7 %, т.е. ≈2/3 поступили на бюджетную форму обучения (Б), 17,1 % – на контрактную (К), 15,2 % – по целевому набору (Ц). Из сравнения средних баллов следует, что наиболее низкие средние баллы как в **ВИ**, так и на **ВК** в такой ФО как К.

Сравнение результатов **ВИ** и **ВК** в зависимости от ФО отражает ситуацию в целом: различия ре-

зультатов **ВИ** и **ВК** высоко значимы в каждой ФО. Аналогичная ситуация наблюдается и, например, в Томском государственном университете систем управления и радиоэлектроники [9].

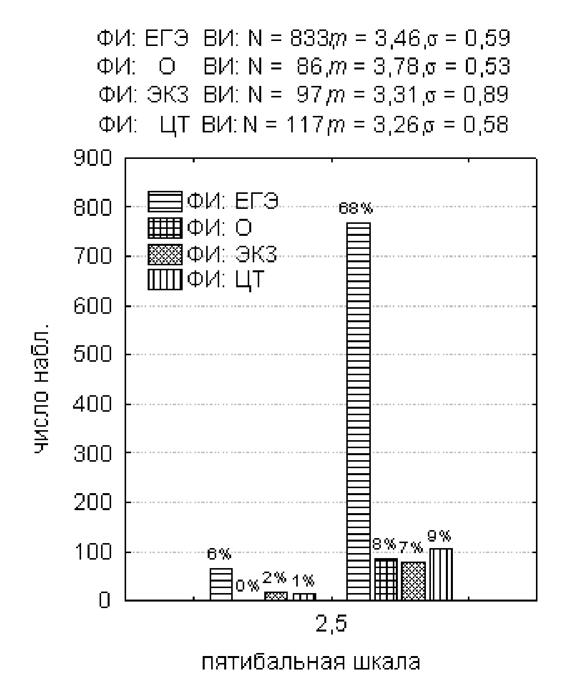


Рис. 6. *Составная гистограмма ФИ* ***ВИ*** *по неравномерной 5-бальной шкале*

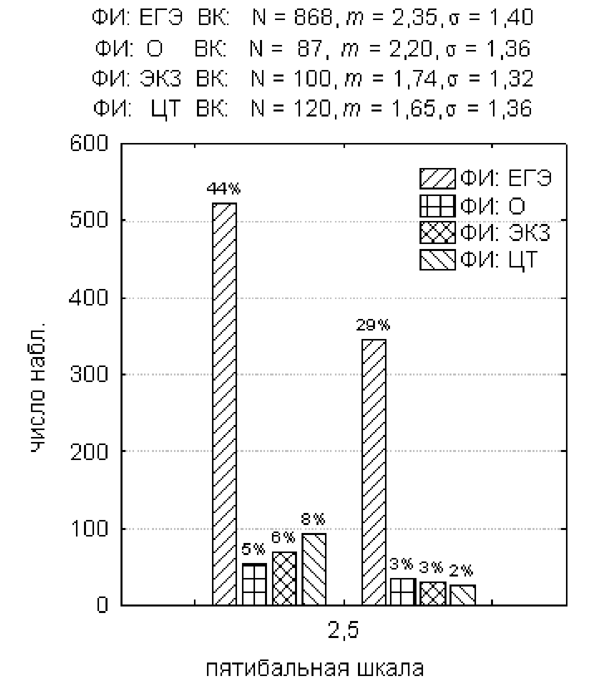


Рис. 7. *Составная гистограмма ФИ* ***ВК*** *по неравномерной 5-бальной шкале*

Представляет также интерес исследование результатов **ВК** в зависимости от диапазонов выборки **ВИ**, отраженное на рис.10, где для простоты восприятия использована неравномерная 5-бальная шкала: [0; 2,5] – «неуд»; (2,5;3,5] – «уд»; (3,5;4,5] – «хор» и (4,5;5] – «отл». Так, например, средний балл в **ВК** студентов выборки **ВИ**:(2,5;3], набравших на **ВИ** от 50 до 60 баллов по 100-бальной шкале, равен 1,581, а средний балл в **ВК** студентов выборки **ВИ**:(3;3,5], набравших в **ВИ** от 60 до 70 баллов по 100-бальной шкале, равен 1,797. Причем, как показывает применение критериев сравнения средних баллов выборок, различие между 1,581 и 1,797 соответствуют уровню значимости $\alpha \approx 0{,}05$, граничному между слабо значимыми и статистически значимыми различиями. Остальные средние баллы двух соседних выборок различаются высоко значимо.

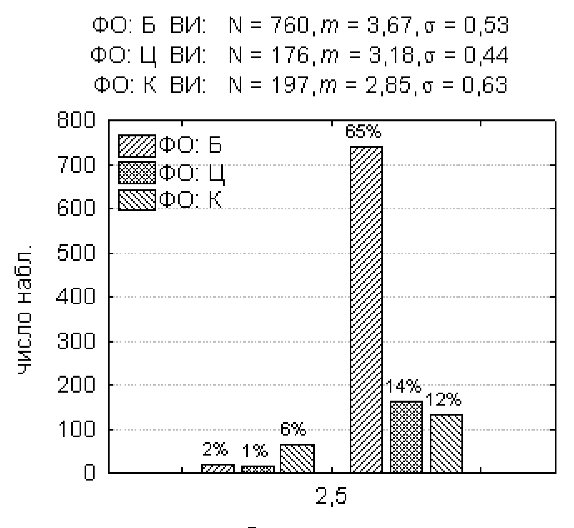


Рис. 8. *Составная гистограмма* ***ФО ВИ*** *по неравномерной 5-бальной шкале*

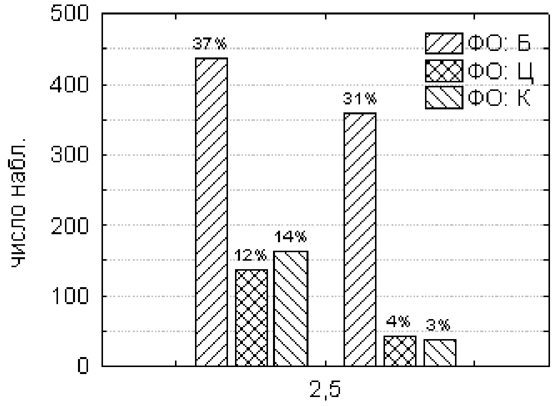


Рис. 9. *Составная гистограмма* ***ФО ВК*** *по неравномерной 5-бальной шкале*

Отметим, что средний балл m выборки **ВК** в табл. 1 представляет собой временной ряд, поддающийся компьютерному анализу и прогнозированию [10]. Применяя метод экспоненциального сглаживания и прогнозирования, можно получить спрогнозированный ряд среднего балла m' выборки **ВК** (табл. 4).

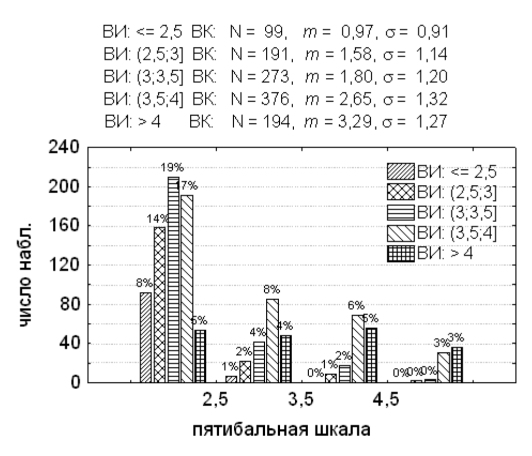


Рис. 10. *Составная гистограмма **ВК** от **ВИ** по неравномерной 5-бальной шкале*

Таблица 4. *Средний балл m и спрогнозированный m′ **ВК** по 5-бальной шкале*

Год	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
m	1,906	1,838	1,788	1,770	2,195	2,250	2,308	2,250	2,381	2,216	–
m′	**1,915**	**1,845**	**1,790**	**1,770**	**2,175**	**2,253**	**2,309**	**2,255**	**2,375**	**2,225**	**2,223**

Примененная версия экспоненциального сглаживания (демпфированного тренда) является высокоточной, и прогнозу на 2005 г. можно было бы доверять, если не учитывать возможную корректировку прогноза ожидаемой «демографической ямой».

Выводы

1. Отличия результатов **ВИ** по математике в ТПУ в 2004 г. (ЕГЭ, централизованное тестирование, олимпиада, вступительный экзамен) от соответствующих результатов **ВК** математических знаний на основе аудиторной контрольной работы с проверкой ее преподавателями являются статистически значимыми.
2. Средний балл **ВИ** в форме ЕГЭ выше среднего балла **ВК** более чем на единицу, что тем более показательно, т.к. из года в год задания по ЕГЭ становятся более сложными и громоздкими (объем полного решения билета ЕГЭ-2004 по математике [11] составляет 15 стр.). Это может говорить об остроте проблемы корректности проведения **ВИ**.
3. Различия в результатах оценки качества знаний по математике методами ТЕСТ и **ВК** являются статистически значимыми.
4. Тестовые механизмы ТЕСТ оценки знаний являются более грубыми. Система тестовых испытаний позволяет получить быстро, но грубо (по системе «зачет») срез знаний студентов в отличие от традиционной системы **ВК** («дифференцированный зачет»).
5. Наблюдаемый рост среднего балла в 1999 году (табл. 1) можно объяснить резким увеличением числа абитуриентов, прошедших обучение на различных подготовительных курсах. В связи с очень низким средним баллом входного контроля для студентов-заочников (≈1,05 по 5-бальной шкале, табл. 2) можно рекомендовать организацию подготовительных курсов на базе представительств и филиалов ТПУ.

СПИСОК ЛИТЕРАТУРЫ

1. Похолков Ю.П. Проблемы и основные направления совершенствования инженерного образования // Alma Mater. Вестник высшей школы. – 2003. – № 10. – С. 3–8.
2. Сенашенко В., Ткач Г. Болонский процесс и качество образования // Alma Mater. Вестник высшей школы. – 2003. – № 8 – С. 8–14.
3. Мотова Г. Система оценки качества образования в странах СНГ и Балтии // Alma Mater. Вестник высшей школы. – 2004. – № 1. – С. 37–40.
4. Финкельштейн В. Уровень математической подготовки выпускников средней школы // Alma Mater. Вестник высшей школы. – 2003. – № 9. – С. 50–51.
5. Боровиков В.П. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. – СПб.: Питер, 2003. – 688 с.
6. Берестова О.Г., Марухина О.В. Компьютерные технологии в оценке качества обучения // Известия Томского политехнического университета. – 2003. – Т. 306. – № 6. – С. 106–112.
7. Марухина О.В., Берестова О.Г. Анализ и обработка информации в задачах оценивания качества обучения студентов вуза // Известия Томского политехнического университета. – 2004. – Т. 307. – № 4. – С. 136–141.
8. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 2002. – 479 с.
9. Иванова Е.В., Рипп А.Г. Пути повышения уровня подготовленности студентов, зачисляемых на первый курс // Современное образование: система и практика обеспечения качества: Матер. регион. научно-метод. конф. – Томск, 2002. – С. 53–54.
10. Боровиков В.П., Боровиков И.П. STATISTICA. Статистический анализ и обработка данных в среде Windows. – М.: Филинъ,1997. – 608 с.
11. Корешкова Т.А., Мирошин В.В., Шевелева Н.В. Математика. Тренировочные тесты ЕГЭ 2004. – М.: ЭКСМО, 2004. – 80 с.